International Journal of Mechanical and Production
Engineering Research and Development (IJMPERD)
ISSN(P): 2249–6890; ISSN(E): 2249–8001
Vol. 10, Issue 1, Feb 2020, 145–156

USE OF REAL-TIME VISUAL IDENTIFICATION TO DESIGN MOVEMENT STRATEGIES FOR INDOOR ROBOTS

FREDY H. MARTÍNEZ. S[1], FERNANDO MARTÍNEZ. S[2] & HOLMAN MONTIEL. A[3]

[1]*Associate Professor, Universidad Distrital Francisco José de Caldas, Facultad Tecnológica, Carrera, Bogotá D.C., Colombia*

[2,3]*Assistant Professor, Universidad Distrital Francisco José de Caldas, Facultad Tecnológica, Carrera, Bogotá D.C., Colombia*

ABSTRACT

Robotics has developed with great vigor in recent years, particularly due to the great interest of researchers, who seek a deeper integration of these artificial systems in the daily life of human beings. One of the areas that is expected to have the greatest impact on society is service robotics. In this area, robotic platforms are developed to perform tasks that support the daily activity of people, for example, caring for children and elderly people. This type of task presents very specific design challenges. For example, the robot must be able to move in the same environment as the human user, however, these types of environments turn out to be dynamic and unknown. The robot must somehow securely identify each of the obstacles in the environment (at least those close by) and define action strategies based on this information. We propose a motion planning scheme for an anthropomorphic robotic platform that relies on the visual identification of specific elements in the environment. From this identification, the strategy defines real-time movement policies that facilitate the programming of tasks in the robot. The strategy was evaluated in the laboratory on our robotic platform, demonstrating high performance and low computational cost.

KEYWORDS: Control Strategy, Image Processing, Motion Planning, Movement Policies, Service Robotics & Visual Identification

Received: Sep 26, 2019; **Accepted**: Oct 16, 2019; **Published**: Dec 24, 2019; **Paper Id.**: IJMPERDFEB202013

Original Article

INTRODUCTION

Human development has been strongly marked by its sense of sight [1]. This is the most important sensor for interacting with the environment and through it, intelligence has been developed, and today, it is a way of transforming the environment through engineering and development. This importance is easily observed in everyday life. For example, when a person, for the first time, is in someplace (cinema, hotel, etc.), his first reaction is to identify visual elements, previously known to trace a strategy of movement in this new environment (look for the ticket office or the reception office). This basic principle of self-localization can be implemented under the same principles on autonomous robots [2].

Great advances in real-time machine vision and image processing schemes have strongly marked the development of autonomous navigation systems for robots. These developments are mainly due to the increased performance of embedded systems and optical sensors with their consequent reduction in cost. At the same time, new image processing schemes have also emerged, including deep neural networks with great possibilities of filtering and classification in real-time [3]. These tools are adapted to the specific needs of the robots and desired

tasks. For example, in the service robotics, tasks are expected to be performed in indoor human environments in unfamiliar areas for the robot (but observable through sensors, optical sensors, for example), and with a high probability of continuous change (dynamic environments) that make programming or prior mapping difficult [4, 5]. These characteristics can also be found in other tasks, in fact, in industrial environments, there are the same problems for the path planning of robots, and just as for service robots, the solution lies in the adoption of reactive strategies from local readings [6, 7]. Under this scheme, the robot defines its movement based on the information collected from the environment, that is, it decides where to move and how to do it if it identifies some important element for the development of its task.

These robots must possess certain processing and control characteristics that allow them to move in the environment. For example, the robot must have a basic function of exploring the environment through which it can sense distinctive elements from shapes, colors, sizes or any other noticeable elements (or combination of them). With this information, the robot can carry out a partial (or even global) reconstruction of the environment without its previous knowledge [8]. This process is usually called self-localization and corresponds to the positioning with respect to some local frame of reference, the landmark identified by the sensors [9]. From this information, the robot can establish distances, relationships and define movement strategies in coherence with its task.

Other types of tasks may have similar characteristics in terms of the problems to be solved. This is the case of Unmanned Aerial Vehicles (UAVs), which independently of their task (surveillance, photo grammetry, supervision, etc.) or control scheme, use similar localization mechanisms. These aerial robots also track specific elements in the terrain below them to find landmarks that coordinate their flight plans. In particular, they first use low-resolution images to find these specific landmarks and then optimize the reading of the images to be used as a reference in the design of the navigation route [10, 11].

It is also possible to use other types of images, beyond those generated by optical transducers based on light, for example, using images produced by ultrasound systems. These images can also be used to process and define navigation strategies and have been used successfully in medicine [12, 13]. Ultrasound systems are capable of producing three-dimensional images (containing 3D tissue information) with high depth detail that can be used to find specific features, and therefore automatically design navigation routes along with tissues for surgical equipment. This considerably reduces possible damage to the patient [14, 15]. This strategy can also be used to plan the route of small robots along the human body [16].

The strategy proposed in this article consists of a general navigation scheme for small service robots with autonomous movement capacity [17, 18]. As a robot platform, we use the Nao robot from Soft Bank Robotics. The optical sensors are the two front cameras located in the head of this robot. The images are filtered for binarization and morphological adjustments to identify specific landmarks in the environment characterized by shape and color [19, 20]. From this identification, the robot executes movement policies that allow it to interact successfully in the environment.

The following part of the paper is arranged in this way. Section 2 presents preliminary concepts and problem formulation. Section 3 illustrates the design profile and development methodology. In Section 4, we present the preliminary results and finally in Section 5, we present our conclusions.

Problem Formulation

Let $W \subset \mathbb{R}^2$ be the closure of a contractible open set in the plane that has a connected open interior with obstacles that represent inaccessible regions. Let Ω be a set of obstacles, in which each $O \subset \Omega$ is closed with a connected piecewise-analytic boundary that is finite in length. Furthermore, the obstacles in Ω are pair wise-disjoint and countably finite in number. Let $E \subset W$ be the free space in the environment, which is the open subset of W with the obstacles removed.

We place an agent (autonomous robot) in the free space of this environment. This agent can know the environment from observations using its sensors. These observations allow it to build an information space I. An information mapping is of the form:

$$q: E \longrightarrow S \tag{1}$$

where,

S denotes an observation space, constructed from sensor readings over time, i.e., through an observation history of the form:

$$\tilde{o}: [0, t] \longrightarrow S \tag{2}$$

The interpretation of this information space, i.e, $I \times S \longrightarrow I$, is that which allows the agent to make decisions. The agent performs readings of the medium through its sensors forming a temporal sequence of values, which is interpreted to perform actions according to the desired movement policy.

We assume the agent is able to sense the proximity, i.e., identify obstacles in the environment using minimal information. The environment E is unknown to the robot. Furthermore, the robot does not even know its own position and orientation. Our goal is to design the control rules for the robot in order to independently solve navigation tasks in a dynamic and unknown environment.

The system is completely independent, i.e., there are no actions on it produced by some superior control unit, internal or external to the robot. The system must actively seek the inherent characteristic of the target and monitor. Trace information is comprised of marks on the navigation environment, landmarks, recognizable by its geometric shape and color. This concept can be extended to any other recognizable trace information.

According to the local information identified by the robot in the environment, a control mode is defined as the determinant of the reactive behavior of the robot. A control mode is a mapping $u: C \longrightarrow \{t, U\}$ that assigns behavior to each possible landmark identified in the environment. These behaviors are defined according to the task of the robot, for example, moving towards the landmark, moving away from it or dodging it. U denotes the set of all possible control modes defined for the robot.

METHODOLOGY

Our recognition scheme uses traditional strategies to identify shapes and colors in images through digital image processing. The overall operation is detailed in the block diagram in figure. 1.

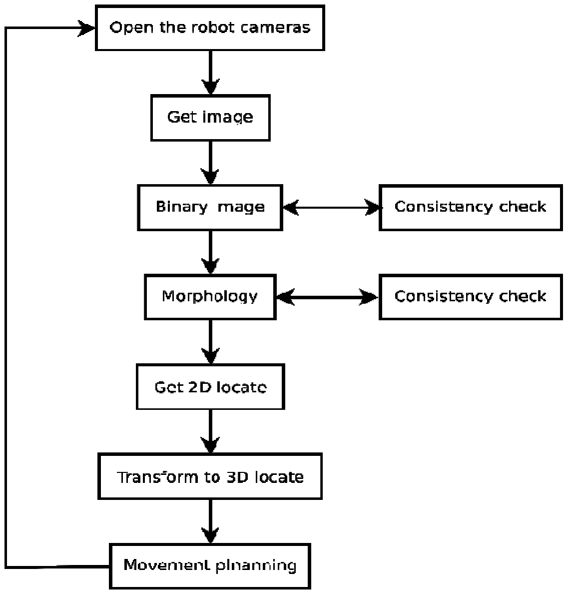

Figure 1: Functional Description of the Identification Algorithm.

Our scheme uses the two cameras of the Nao robot (top and bottom). The code is implemented in the Nao robot using Naoqi. The video from the cameras is captured at 15 frames per second in RGB color model (color model in which Red, Green and Blue light are added together) with a frame size of 640*480 pixels (kVGA resolution). The frames are not scaled, all image processing is done in the same resolution. All images are captured and stored in PNG (Portable Network Graphics) format.

The first filter applied to each frame is the binarization of the image in two colors. This binarization is done with Open CV in the HSV color space (Hue, Saturation, Value; alternative representations of the RGB color model) using as pattern a color involving yellow, red, or blue. Then, we perform morphological image processing on the images to identify basic geometric shapes. The initial tests have been developed with yellow circles.

Once the regions of interest have been identified, they are labeled and characterized. Using Numpy matrix operations, the 2D location of the object in the image is identified. For verification, this information is superimposed on the original image captured by the Nao's camera (the region is marked with a circle and its center, from which the three-dimensional coordinates are also defined). Then, we transform the 2D location to an absolute distance using the principle of ranging. The estimation is not completely accurate due to the lack of information regarding the depth; however, combining the information from the two cameras achieves a value quite close to the real. Finally, the Nao robot is programmed to respond in coherence with the identified object (walk to the ball). This last step consists of transferring the estimated 3D measurements from the images into a 3D location system on the environment, which allows the definition of movement policies to the robot’s joints.

All our searches and recognition schemes are written in Python 3.7.3 with the use of Open CV 4.1.0.25, Numpy 1.16.2, Pillow 5.4.1 and Naoqi. Figure 2 shows the result of one of the laboratory tests (object to recognize: yellow ball).

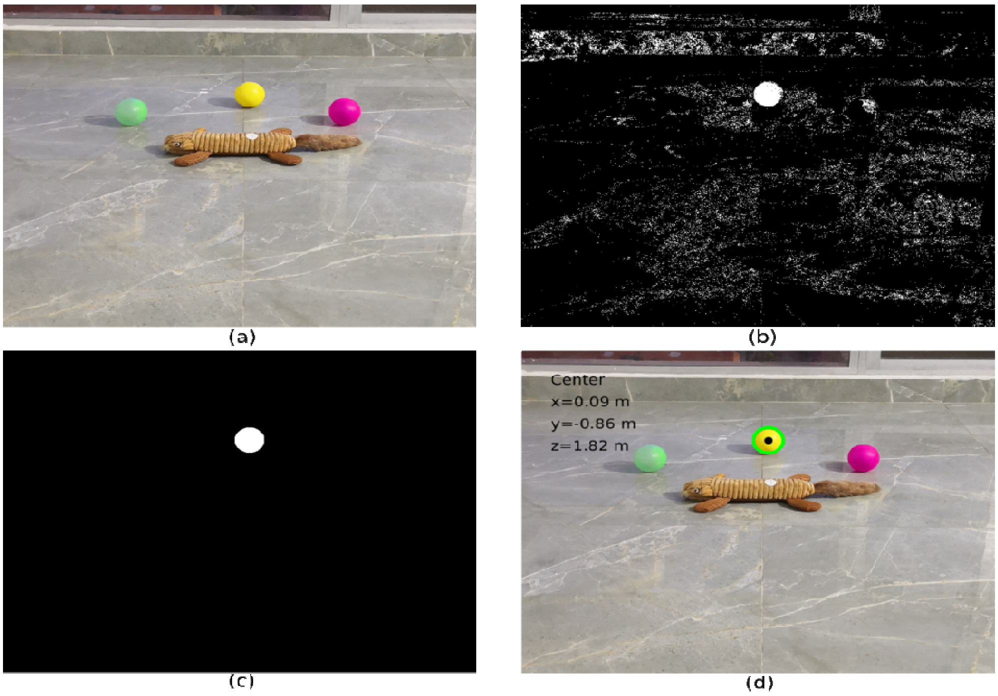

Figure 2: Operation of the Algorithm in the Laboratory. (a) The Three Balls Used for Evaluation: Yellow, Green and Red. (b) Image Binarized by HSV Color Space. (c) Identification of Regions by Morphological Adjustment and (d) Initial Image with Superimposed Localization Information.

RESULTS AND DISCUSSIONS

The tests were performed on our robotic platform. Our assistive robot consists of two robotic platforms: A humanoid Nao robot from Soft Bank Group for interaction with humans and the environment and an ARMOS Turtle Bot 1 robot from the ARMOS research group for indoor navigation (Figure 3). Communication with the two platforms is via a Wi-Fi connection. The Nao robot is used for image capture, real-time processing and as an interface with the environment (interaction with arms and head movement), while the ARMOS Turtle Bot 1 robot is used for platform navigation in the environment. In our initial tests, the global control of the scheme is performed by the Nao robot, which informs the ARMOS Turtle Bot 1 robot how to move in the environment; however, the future development contemplates performing the processing and control on the control card of the ARMOS Turtle Bot 1 robot.

We evaluate the performance of the strategy in the laboratory with different configurations varying the position of the balls, distances to the robot, number of balls and even different lighting conditions. Despite the great possibilities offered by the environment, the algorithm was always able to correctly identify the object of interest. In some frames, the algorithm confused the ball with the environment when the light conditions were particularly poor; however, from neighboring images, it was possible to establish the 2D location of the robot in 100% of cases.

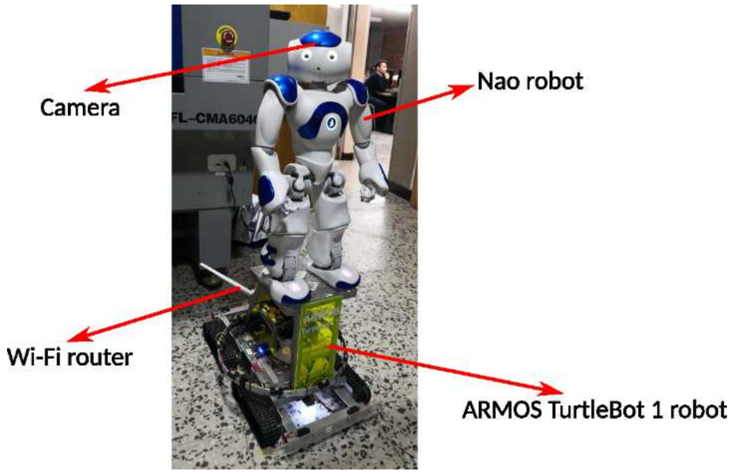

Figure 3: Experimental Setup for the Identification System. It is Composed of a Humanoid Nao Robot from Soft Bank Group at the Top and an ARMOS Turtle Bot 1 Tank Robot from the ARMOS Research Group at the Bottom.

To define the relative position of the object of interest or landmark to the robot, we define a three-dimensional rectangular coordinate system with origin in the robot camera (Figure. 4). The *x*-axis grows positively to the right of the robot, the *z*-axis grows positively to the front of the robot, and the *y*-axis grows positively upwards from the robot head. This means that elements on the ground will have a negative component in the *y*-axis (they are below the level of the head, Figure. 2d). The camera has a viewing angle of 70°, which together with the size of the robot defines a detection field in front of the robot for the definition of reactive movements.

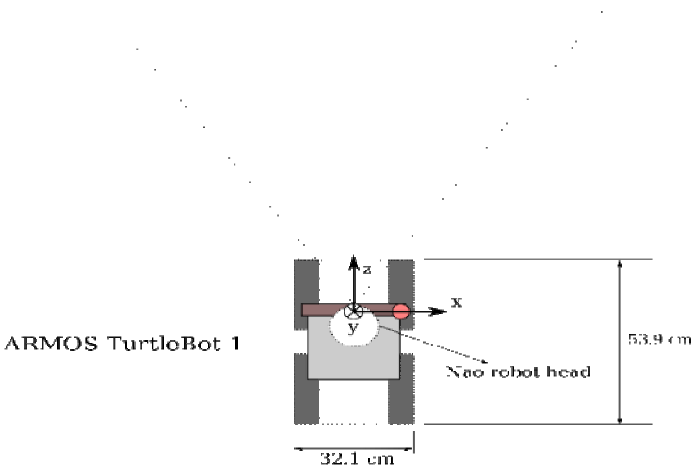

Figure 4: Three-Dimensional Reference System with Origin in the Optical Sensor of the Robot used to Establish the Relative Position of the Landmark to the Robot.

Figure 2 shows the results for a performance test under controlled conditions. We perform many of these tests to evaluate the performance of the scheme. Table 1 shows a summary of 20 of these tests performed in the same environment under the same lighting conditions. In these tests, we change the location of the target object to different points in the environment, all within the robot's field of vision. The code used was always the same, identifying the yellow ball. The identification of the landmark was achieved in 100% of the cases. Also, the estimation error to the actual distance to the landmark averaged less than 3%. A slight increase in the error is observed when the distances increase, but the error value never exceeds 7% (Figure 5).

Table 1: Outline of 20 Laboratory Tests in which the Position of the Landmark is Estimated in Different Situations, all within the Robot's Field of Vision

Estimated three-dimensional coordinates			Estimated straight line distance [m]	Actual distance to sensor [m]	Estimation error [%]
x [m]	y [m]	z [m]			
0.09	-0.86	1.82	2.01	1.96	2.85
-0.15	-0.75	1.51	1.69	1.63	4.09
-0.36	-0.79	1.93	2.12	1.98	6.66
0.12	-0.91	2.03	2.23	2.17	2.82
0.26	-0.49	0.56	0.79	0.77	2.42
0.39	-0.56	0.76	1.02	1.00	1.78
-0.43	-0.78	0.83	1.22	1.19	2.05
-0.05	-0.76	0.79	1.10	1.09	0.24
-0.18	-0.81	0.86	1.20	1.15	4.08
0.15	-0.78	1.33	1.55	1.52	1.89
0.38	-0.67	1.56	1.74	1.65	5.39
0.49	-0.69	1.79	1.98	1.93	2.80
0.21	-0.40	0.69	0.82	0.79	3.55
0.08	-0.43	0.46	0.63	0.63	0.72
0.26	-0.35	0.28	0.52	0.52	0.38
0.36	-0.48	0.81	1.01	0.99	2.31
0.42	-0.46	0.68	0.92	0.89	3.38
-0.13	-0.37	0.73	0.83	0.81	2.02
-0.26	-0.45	0.76	0.92	0.89	3.56
0.38	-0.42	0.97	1.12	1.10	2.46
			Average error:		**2.77**

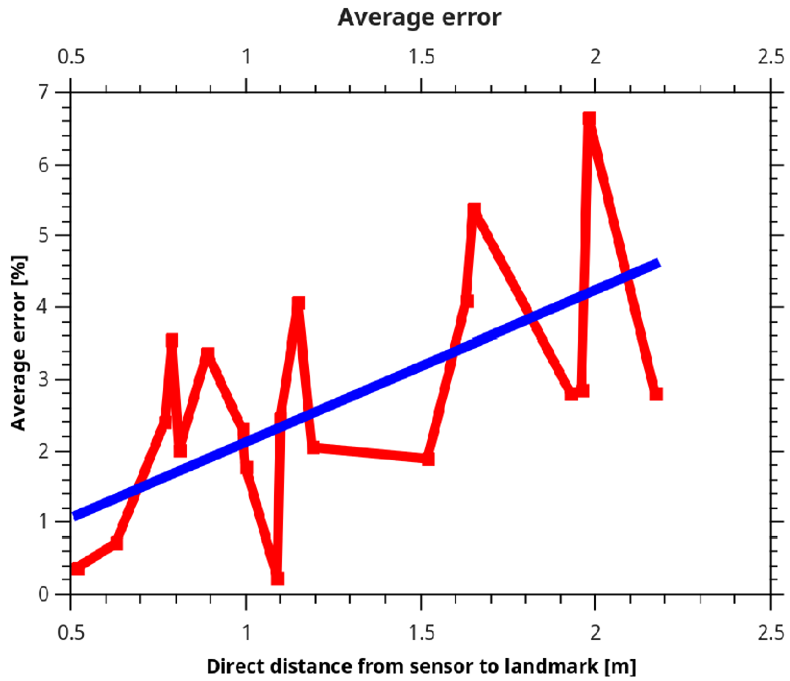

Figure 5: Error Behavior Concerning the Distance Between the Optical Sensor and the Landmark.

The reactions programmed in the robot as a response to the identification of a landmark correspond to the basic movements of the robot in the environment coherent with its task. In principle, we develop the task of picking up a ball and carrying it to the trash can, so the control modes programmed as a response include: walking to the landmark, moving away from the landmark and dodging the landmark. These tasks were scheduled parallel to our identification and tracking scheme and presented an excellent performance in real-time (without delays that would affect the task's performance).

From the results, it is proposed to improve the algorithm by including stereoscopic vision. In our platform, we have the problem of incorporating a system of two cameras to the robot, or in its defect, to add some sensor that is able to inform about the depth to the object of interest.

CONCLUSIONS

To develop a protocol for the development of applications for an assistive robot, we present in this article a strategy for the identification of characteristics in the environment, from which it is possible to define the development of parallel tasks and movement control policies. The algorithm uses Open CV to identify the elements of interest from colors and shapes.

In particular, we have evaluated the operation by filtering through yellow, blue and red colors, and for circular shapes. The tests were performed with balls of different colors within reach of the robot's cameras. The scheme uses color binarization and morphological adjustment over the regions to determine the target point. Once the area has been identified in the 2D image, this information is tagged and transformed into 3D location to coordinate the robot's movement.

The performance tests were performed on a Nao V5 Evolution robot equipped with an Intel Atom @ 1.6 GHz processor and OS NAOqi 2.0. The cameras on the head of the Nao robot were used as optical sensors in the strategy. The code was developed in Python and Open CV Laboratory tests showed high algorithm performance and very low computational cost. The error in the estimation of the distance from the sensor to the target never exceeded 7% with respect to the actual distance. In addition, the error is reduced as the distance to the target decreases. Besides, integration with parallel tasks could be carried out without causing operational delays.

In order to reduce the error in the estimation, we propose to integrate in the future to the ARMOS Turtle Bot 1 robot a scheme of two cameras with a disposition similar to the one of the eyes in people and animals, this with the purpose of adding information to the scheme related to the depth at which the landmark is located. This goes hand in hand with the migration of the control scheme to a higher capacity processing unit located in this robot.

ACKNOWLEDGMENTS

This work was supported by Universidad Distrital Francisco José de Caldas and the Centre for Scientific Research and Development (CIDC) through the project 1-72-578-18. The views expressed in this paper are not necessarily endorsed by Universidad Distrital Francisco José de Caldas or the CIDC. The authors thank the research groups ARMOS and SIEand its research seedbeds for the evaluation carried out on prototypes of ideas and strategies proposed in this paper.

REFERENCES

1. *Kaur, B. and Bhattacharya, J. (2016). Predictive hierarchical human augmented map generation for itinerary perception. Electron. Lett., 52(16):1381–1383, ISSN 0013-5194, doi:10.1049/el.2016.0397.*
2. *Bista, S., Giordano, P., and Chaumette, F. (2016). Appearance-based indoor navigation by IBVS using line segments. IEEE Robot. Autom. Lett., 1(1):423–430, ISSN 2377-3766, doi:10.1109/LRA.2016.2521907.*
3. *Jose, J., Dinakaran, D., Ramya, M., & Samuel, D. H. A Survey on Magnetic Wall-Climbing Robots for Inspection.*
4. *Faisal, M., Mathkour, H. and Alsulaiman, M. (2018). I3ms: Intelligent multi-sensor multi-baseline mapping system. IEEE Access, 6(1):4243–4254, ISSN 2169-3536, doi:10.1109/ACCESS.2017.2788422.*

5. *Durand-Petiteville, A., Le Flecher, E., Cadenat, V., Sentenac, T. and Vougioukas, S. (2018). Tree detection with low-cost three-dimensional sensors for autonomous navigation in orchards. IEEE Robot. Automat. Lett., 3(4):3876–3883, ISSN 2377-3766, doi:10.1109/LRA.2018.2857005.*
6. *Sundari, P. M., & Kumar, S. B. R. (2014). A study of image processing in analyzing tree ring structure. Int. J. Res. Humanit. Arts Lit, 2(3), 13-18.*
7. *Rendón A. (2015). Evaluación de estrategia de navegaciónautónomabasadaencomportamientoreactivo para plataformasrobóticasmóviles. Tekhnê, 12(2):75–82, ISSN 1692-8407.*
8. *Nakhaeinia, D., Payeur, P., and Laganiere, R. (2018). A mode-switching motion control system for reactive interaction and surface following using industrial robots. IEEE/CAA J. Autom. Sin., 5(3):670–682, ISSN 2329-9266, doi:10.1109/JAS.2018.7511069.*
9. *Monica, R. and Aleotti, J. (2018). Surfel-based next best view planning. IEEE Robot. Autom. Lett., 3(4):3324–3331, ISSN 2377-3766, doi:10.1109/LRA.2018.2852778.*
10. *Meng, Z., Qin, H., Chen, Z., Chen, X., Sun, H., Lin, F. and Jr. M. (2017). A two-stage optimized next-view planning framework for 3-D unknown environment exploration and structural reconstruction. IEEE Robot.. Autom. Lett., 2(3):1680–1687, ISSN 2377–3766, doi:10.1109/LRA.2017.2655144.*
11. *Lv, W., Kang, Y. and Zhao, Y. (2019). Fvc: A novel nonmagnetic compass. IEEE Trans. Ind. Electron., 66(10):7810–7820, ISSN 0278-0046, doi:10.1109/TIE.2018.2884231.*
12. *Dudhrejia, M. N., Bhatt, C. B., & Shah, M. K. Unexplored Idea to Examine Grain Specimen Quality by Utilizing Image Processing Intelligence.*
13. *Liu, C., Zhang, S. and Akbar, A. (2019). Ground feature oriented path planning for unmanned aerial vehicle mapping. IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens., 12(4):1175–1187, ISSN 1939-1404, doi:10.1109/JSTARS.2019.2899369.*
14. *Li, J., Deng, G., Luo, C., Lin, Q., Yan, Q. and Ming Z. (2016). A hybrid path planning method in unmanned air/ground vehicle (uav/ugv) cooperative systems. IEEE Trans. Veh. Technol., 65(12):9585–9596, ISSN 0018–9545, doi:10.1109/TVT.2016.2623666.*
15. *Hennersperger, C., Fuerst, B., Virga, S., Zettinig, O., Frisch, B., Neff, T. and Navab, N. (2017). Towards MRI-based autonomous robotic us acquisitions: A first feasibility study. IEEE Trans. Med. Imag., 36(2):538–548, ISSN 0278-0062, doi:10.1109/TMI.2016.2620723.*
16. *Pattni, V. B., Naveenchandran, P., Thamotharan, C., & Rajasekar, R. Hcci Combustion: Mathematical Modelling Approach Using Visual Basic for Applications.*
17. *Huang, Q., Lan, J. and Li X. (2019). Robotic arm-based automatic ultrasound scanning for three-dimensional imaging. IEEE Trans. Ind. Informat., 15(2):1173–1182, ISSN 1551-3203, doi:10.1109/TII.2018.2871864.*
18. *Lu, B., Chu, H., Huang, K. and Cheng, L. (2019). Vision-based surgical suture looping through trajectory planning for wound suturing. IEEE Trans. Autom. Sci. Eng., 16(2):542–556, ISSN 1545-5955, doi:10.1109/TASE.2018.2840532.*
19. *Huang, Q., Wu, B., Lan, J. and Li X. (2018). Fully automatic three-dimensional ultrasound imaging based on conventional B-scan. IEEE Trans. Biomed. Circ. Syst. , 12(2):426–436, ISSN 1932-4545, doi:10.1109/TBCAS.2017.2782815.*
20. *Vandini, A., Bergeles, C., Glocker, B., Giataganas P. and Yang, G. (2017). Unified tracking and shape estimation for concentric tube robots. IEEE Trans. Robot., 33(4):901–915, ISSN 1552-3098, doi:10.1109/TRO.2017.2690977.*

21. *McGinn, C. (2019). Why Do Robots Need a Head? The Role of Social Interfaces on Service Robots. Int. J. Soc. Robot. . 11(49):1–15, ISSN 1875-4791, doi: https://doi.org/10.1007/s12369-019-00564-5.*

22. *Zhang, S., Qin, J., Cao, S. and Dou, J. (2018). HRI Design Research for Intelligent Household Service Robots: Teler as a Case Study Design, User Experience and Usability: Designing Interactions. 10919(1):513–524, ISSN 0302-9743, doi: https://doi.org/10.1007/978-3-319-91803-7_39.*

23. *Castañeda, J. and Salguero, Y. (2017). Adjustment of visual identification algorithm for use in stand-alone robot navigation applications. Tekhnê. 14(1):73–86, ISSN 1692-8407.*

24. *Sanchez, J., Wang, M., Olivares, M., Molina, M. and Voos, H. (2019). A Real-Time 3D Path Planning Solution for Collision-Free Navigation of Multirotor Aerial Robots in Dynamic Environments. J. Intell. Robot. Syst. 93(1–2):33–53, ISSN 1573-0409, doi: https://doi.org/10.1007/s10846-018-0809-5.*

AUTHORS PROFILE

Fredy H. Martínez S. is a associate professor at the Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Bogotá D.C.-Colombia. He obtained his Bachelor's Degree in Electrical Engineering and his Ph.D in Engineering - Systems and Computing from the National University of Colombia (Colombia) in 1997 and 2018 respectively. Since 2000 he leads the ARMOS research group at the Universidad Distrital Francisco José de Caldas (Colombia). His research focuses on electronics, control systems, hybrid architectures, autonomous robotics and intelligent systems. The application of robotic systems in the provision of services to people has recently been addressed.

Fernando Martínez S. is a assistant professor at the Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Bogotá D.C.-Colombia. He obtained his Bachelor's Degree in Control Engineering and his Master Degree in Electronic and Computer Engineering from the Universidad Distrital Francisco José de Caldas (Colombia) in 2004 and

2012 respectively. His research focuses on the fields of electronics, instrumentation systems, real-time image and video processing, embedded signal processing solutions. Recently, the development of autonomous navigation strategies based on images has been tackled.

Holman Montiel A. is a assistant professor at the Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Bogotá D.C.-Colombia. He obtained his Bachelor's Degree in Control Engineering and his Master Degree in Computer Security from the Universidad Distrital Francisco José de Caldas (Colombia) in 2005 and 2014 respectively. His field of research focuses on signal processing, embedded systems, and custom-designed hardware and software cryptography schemes.